2-2007

Model-Based Estimation of Muscle Forces Exerted During Movements

Ahmet Erdemir
Cleveland Clinic Foundation

Scott McLean
Cleveland Clinic Foundation

Walter Herzog
University of Calgary

Antonie J. van den Bogert
Cleveland State University, a.vandenbogert@csuohio.edu

Follow this and additional works at: http://engagedscholarship.csuohio.edu/enme_facpub

Part of the Biomechanical Engineering Commons

How does access to this work benefit you? Let us know!

Publisher's Statement
NOTICE: this is the author's version of a work that was accepted for publication in Clinical Biomechanics. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Clinical Biomechanics, 22, 2, (02-01-2007); 10.1016/j.clinbiomech.2006.09.005

Original Citation

This Article is brought to you for free and open access by the Mechanical Engineering Department at EngagedScholarship@CSU. It has been accepted for inclusion in Mechanical Engineering Faculty Publications by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.
Model-based estimation of muscle forces exerted during movements

Ahmet Erdemir a, Scott McLean a, Walter Herzog b, Antonie J. van den Bogert a,*

a Department of Biomedical Engineering (ND-20), The Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA
b Human Performance Laboratory, The University of Calgary, Calgary, AL T2N 1N4, Canada

1. Introduction

The force output of cardiac muscle can be quantified by simply recording arterial pressure. It is far more difficult to obtain clinically relevant information on the function of skeletal muscles. Imagine what could be done with such information. In the treatment of cerebral palsy, the clinician could “see” which muscle is responsible for an abnormal gait pattern, and that muscle could then be directly targeted for surgery. In an athlete with a recurrent overuse injury, we could “see” the loads being placed upon bones and joints during movement and how these loads are altered during rehabilitation. There are many other neurological and orthopaedic problems where knowledge of muscle forces could enhance clinical decision making. In this paper, we will review what methods currently exist, and to what extent they are ready for clinical applications.

Direct measurement of muscle forces in vivo is usually limited to minimally invasive measurements in superficial tendons such as the Achilles (Finni et al., 1998; Komi et al., 1992). Otherwise, in vivo measurements can be conducted in the operation room where a force transducer can be placed on a tendon, following data collection and the removal of the device before the completion of the surgery, e.g. flexor tendons of fingers during surgeries of carpal tunnel (Dennerlein et al., 1998; Dennerlein et al., 1999; Dennerlein, 2005; Schuind et al., 1992). Such approaches may not necessarily be feasible in a clinical setting; therefore such tendon force measurement techniques have been utilized mostly in research laboratories (Ravary et al., 2004; Fleming and Beynnon, 2004). Non-invasive methods rely
on the basic principle that muscles produce skeletal movement and ground reaction forces. Clearly, none of these observable variables provides information on any single muscle. Instead, a technique known as inverse dynamic analysis has been developed, based on computational modeling of the dynamics of linked body segments. The analysis produces estimates of joint torques, each of which represents the resultant action of all muscles crossing a joint. While inverse dynamic analysis has become a routine tool in clinical gait analysis (Vaughan et al., 1992; Winter, 2005), muscles are not represented and the approach provides no information on muscular load sharing, agonist–antagonist activity, energy transfer between joints via biarticular muscles, and dynamic coupling (van den Bogert, 1994; Zajac et al., 2002). Electromyography (EMG) data can support a clinical inverse dynamic analysis to more effectively interpret joint torques, but there are no estimates of individual muscle forces (Zajac et al., 2003).

Actual estimates of muscle forces can only be obtained with computational models in which the skeleton and muscles are both represented. Implemented in a variety of forms, musculoskeletal models have been used in conjunction with non-invasive measurements to obtain individual muscle forces during a number of movement tasks. Within the current article, we have attempted to critically evaluate those studies that have combined musculoskeletal models, optimization methods and movement data to estimate individual muscle forces. A review of literature is first provided with the necessary methodological background, followed by the applications of the various techniques with a discussion of limitations. Novel strategies that attempt to improve understanding of muscle function are also presented. We will conclude with recommendations, for clinical applications and for further research that may increase the applicability and validity of these techniques in clinical practice.

2. Musculoskeletal dynamics

Dynamic human motion is achieved via activation of the muscles, which subsequently produce force and in turn, move the joints in a controlled fashion to accomplish the predetermined task requirements. Quite often, these tasks are also required to take place against the action of external forces. The outcome of this entire process largely depends on the force-generation properties of the muscles, the anatomical features of the skeletal system (e.g. anthropometric properties, muscle paths) and the underlying neuronal control system. It is thus critical to understand the coupling between these mechanisms if one wishes to examine the success and applicability of various muscle force estimation techniques. A brief explanation on modeling these various components is thus presented below.

2.1. Equations of motion

For illustrative purposes, we will consider a musculoskeletal system where the kinematic degrees of freedom (DoF) are a set of \( n \) joint angles \( q \). The relationship between movement and muscle forces in a musculoskeletal model (Fig. 1A) can be expressed in matrix form by Eq. (1) (Pandy, 2001).

\[
M(q)\ddot{q} + C(q, \dot{q}) + G(q) + R(q)F_{MT} + E = 0, \quad (1)
\]

where \( M(q) \) is the system mass matrix \((n \times n)\); \( C(q, \dot{q}) \) is the centrifugal and coriolis loading \((n \times 1)\); \( G(q) \) is the gravitational loading \((n \times 1)\); and \( E \) represents external forces. \( R(q)F_{MT} \) represents muscular joint torques \((n \times 1)\), where \( R(q) \) is the matrix of muscular moment arms \((n \times m)\) and \( F_{MT} \) are the muscle forces \((m \times 1; m: \text{number of muscles})\).

The system is usually redundant, with the number of unknown muscle forces exceeding the number of equations \((m > n)\). In order to estimate muscle forces therefore, one must either reduce \( m \) by combining muscles (Pierrynowski and Morrison, 1985); or use a methodology relying on optimization principles (described below). In its simplest form, the reduction approach converges to the standard gait analysis protocol where the muscular torque at each joint is calculated from movement data and the ground reaction forces (Otten, 2003; Fig. 1B). In that case, the generalized system equations reduce to allow one-to-one correspondence between degrees of freedoms and muscular loading:

\[
M(q)\ddot{q} + C(q, \dot{q}) + G(q) + T_{MT} + E = 0, \quad (2)
\]

where \( T_{MT} \) are the muscular joint torques \((n \times 1)\) which are equal to \( R(q)F_{MT} \).

2.2. Muscle–skeleton coupling

The origin and insertion sites of the muscles of interest define the associated moment arms at the joints that they span. This moment arm, multiplied by the force generated

![Fig. 1. (A) A musculoskeletal model of the lower extremity utilized in forward dynamics solutions. Muscle forces generate the movement of the hip, knee, and ankle joints. (B) A joint torque-driven model of the lower extremity commonly used for inverse dynamics analysis of gait data. Lower extremity was simply illustrated as an example; similar models exist for upper extremities and other body segments.](image-url)
by the muscle, is the magnitude of the muscle force contribution to the resultant joint torque which generates joint rotation. Moment arms are classically defined as the distance between a muscle’s line of action and the joint’s axis of rotation. As indicated in Eq. (1), this distance may be joint angle dependent.

A more general definition of moment arms is provided using the principle of virtual work (e.g., An et al. (1984b)). If \( R_j(q) \) for example, is the moment arm of muscle \( j \) with respect to joint axis \( i \), the following equation holds:

\[
R_j(q) = -\frac{\partial L_j(q)}{\partial q_i},
\]

where \( L_j(q) \) is the origin–insertion length of muscle \( j \) as a function of all joint angles \( q \). Considering this relationship, a musculoskeletal model that is used for estimating muscle forces must necessarily incorporate accurate and anatomical descriptions of the muscle insertions and the three-dimensional path of the muscle relative to the moving skeleton (Delp and Loan, 1995; Herzog, 1992).

2.3. Muscle modeling

The magnitude of the muscle force depends on its activation level (from the activation dynamics, Eq. (4a)) as well as its force-generation properties defined by force–fiber length and force–fiber velocity relationships (Eq. (4b), Fig. 2). Also important are the properties of the tendon, which is serially attached to the muscle and completes the musculotendon unit (Zajac, 1989). Zajac (1989) presented the widely known Hill-type muscle model in a generic fashion as two differential equations:

\[
\dot{a} = f_1(u, a),
\]

\[
\dot{l}_M = f_2(l_M, l_{MT}, a),
\]

where \( u \) is the muscle excitation and \( a \) is the muscle activation. \( l_M \) is the muscle fiber length and \( l_{MT} \) is the musculotendon unit length. Muscle force (\( F_{MT} \)) is the by-product of the solution of these dynamic equations. It is common practice to represent any musculotendon unit by defining the following model parameters: maximum isometric force (\( F_0 \)), optimal fiber length (\( l_{0M} \)), maximum shortening velocity (\( v_S \)), tendon slack length (\( l_{TS} \)) and pennation angle (\( \alpha \)) (Zajac, 1989). For a given state, all these intrinsic muscle properties influence the magnitude of the muscle force, and therefore define the boundaries of maximal muscle force during muscle force estimations. This model has been widely accepted and used in many large-scale musculoskeletal models as well as commercially available software such as SIMM (Musculographics, Inc., Chicago, IL, USA; Delp and Loan (1995)).

2.4. Forward solution

The system equations (Eq. (1) or Eq. (2)) provides the relationship between skeletal motion and muscle forces or joint torques occurring during the movement under investigation. How this equation is used depends on the research/clinical problem and the availability of the experimental data or a priori information related to the movement. When muscle excitations or joint torques are available or assumed, a forward dynamics approach can be utilized that integrates the system equations to calculate the movement patterns (Fig. 3A):

\[
\ddot{q} = \frac{1}{M(q)}[C(q, \dot{q}) + G(q) + T_{MT} + E].
\]

The approach is advantageous in that the movement is predicted. Yet, accurate (a priori) knowledge of muscle excitations (forces) or joint torques is rare, eliminating the stand-alone application of this technique. This method, however, can be utilized in combination with optimization to estimate muscle forces. Examples of this will be reviewed below. In forward solutions, \( E \) is usually obtained using a viscoelastic contact model (\( E(q, \dot{q}) \); e.g., McLean et al., 2003). A less common alternative is to use measured external loads as input (\( E(t) \)) but this can lead to unstable solutions.

2.5. Inverse solution

The inverse dynamics approach has been a frequent component of muscle force estimation routines, due to the availability of the joint kinematics data and ground reaction forces following a standard gait analysis. Given the time history of these variables, it is possible to calculate

![Fig. 2. Commonly used musculotendon model for musculoskeletal simulations. The contractile element (CE) of the muscle is in parallel with the passive element (PE); all in series with the tendon. Force generation capacity of the muscle was defined by the force–length and force–velocity properties of the contractile element and the nonlinear spring properties of the passive element. In the most general form, tendon elasticity is assumed to be nonlinear and pennation angle (\( \alpha \)) is included in the calculations.](image-url)
muscular joint torques at each instant of the movement by re-arranging Eq. (2) (Fig. 3B):

\[ T_{MT} = M(q)q + C(q, \dot{q}) + G(q) + E. \]  

(6)

Although Eq. (6) is a straightforward representation of the system dynamics to solve for muscular joint torques, it is rarely used in practice. Instead, muscular joint torques are typically derived from equations of motion of a single segment, working recursively from distal to proximal (Winter, 2005). A notable exception is the work by Kuo (1998), who applied Eq. (6) to a whole body model. Such an approach was found to be advantageous since there are typically fewer unknown joint torques than equations (degrees of freedom) in a whole body model. \( q \) includes translational and rotational degrees of freedom for the body relative to the ground where the forces and moments are known to be zero.

The inverse dynamics technique has also been utilized to evaluate gait changes as a result of pathology or treatment (Davids et al., 2004). These types of investigations are descriptive, not predictive, and interpretation at the muscular level is necessarily based on a total muscular joint torque, and also possibly on EMG data. With the help of optimization techniques described in the following sections, the methodology can be used to estimate muscle forces, therefore providing an in-depth evaluation of muscle function during the measured movement.

3. Muscle force estimation

Model-based estimation of muscle forces usually requires optimization regardless of the strategy (inverse or forward dynamics) selected to solve for the equations describing the musculoskeletal system (Pandy, 2001; Tsirakos et al., 1997). The redundancy of muscular load sharing can be addressed by minimizing a cost or objective function appropriately selected for the movement under investigation. The adoption of either an inverse or forward dynamics approach is typically dependent on the availability of the experimental data or the clinical/research question to be answered.

3.1. Inverse dynamics-based static optimization

Muscle force estimation using gait data combined with inverse dynamics and static optimization has been practiced for almost three decades (Tsirakos et al., 1997). First, joint torques are calculated from joint kinematics and ground reaction force data using Eq. (6). The muscular load sharing problem is then solved for each instant in time, by minimizing an objective function \( J \) (e.g. total muscle force) subject to constraints representing the equality of the sum of individual muscular moments to the joint torques calculated from the inverse dynamics analysis (Fig. 4A, Eq. (7)). The individual muscular moment is calculated from the muscle force (the unknown of the optimization problem) and muscle moment arms, which are derived from musculoskeletal anatomy and may or may not depend on joint angles. Usually, the maximum possible muscle forces are limited by physiological values as an additional boundary constraint (Eq. (7)). Muscular dynamics can be implicitly implemented by deriving time-dependent bound constraints on muscle force from lower and upper bounds of excitation levels fed through a dynamic muscle model (Happee, 1994; Happee and van der Helm, 1995). The optimization problem may be subject to additional constraints \( (g,h) \) depending on the specifics of the joint under investigation, e.g. constraints on the direction of joint contact force to prevent dislocation of the glenoid joint during simulations (van der Helm, 1994).

minimize \[ J(F_{MT}) \]

subject to \[ R(q)F_{MT} = T_{MT}, \]
\[ 0 \leq F_{MT} \leq F_{max}, \]
\[ g(F_{MT}, q) \leq 0, \]
\[ h(F_{MT}, q) = 0. \]  

(7)
Inverse dynamics-based static optimization has been commonly applied to estimate muscle forces in the lower extremity during walking (Table 1). Several investigators tested the sensitivity of the results on objective function selection (e.g., Glitsch and Baumann, 1997; Collins, 1995), as well as model parameters, (e.g. Glitsch and Baumann, 1997; Herzog, 1992; Brand et al., 1986). Many simpler models and movements were also investigated for these purposes (Table 1). Static optimization is computationally efficient since it does not require multiple integrations (solution of Eq. (6) instead of Eq. (5)). The moment equality and boundary constraints of muscle forces are typically linear and proper selection of an objective function can further increase cost-effectiveness by reducing the optimization problem into linear programming, e.g. total muscle force (Rohrle et al., 1984; Pedersen et al., 1987). Although limited, the linear programming can be implemented into standard movement analysis in a straightforward fashion. In particular, a double linear programming formulation (spinal compression force subject to minimized maximum muscle intensity) was frequently used for the analysis of spinal forces (Table 1) and implemented
<table>
<thead>
<tr>
<th>Activity</th>
<th>Model</th>
<th>Objective</th>
<th>Validation</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking</td>
<td>31 MGs, 7 Segments</td>
<td>Minimize sum of muscle forces + 4 * (sum of joint moments) and also original + weighted hip joint force</td>
<td>EMG, comparison of hip joint forces with literature</td>
<td>Predicted muscle load sharing and lower limb joint reactions. Modifying objective function to include minimization of joint reactions resulted in a limited reduction in joint forces</td>
<td>Seireg and Arvikar (1975)</td>
</tr>
<tr>
<td></td>
<td>27 MGs, 4 Segments</td>
<td>Minimize sum of muscle stresses</td>
<td>Temporal validation of muscle force predictions via EMG</td>
<td>Accurate determination of hip joint contact and muscle forces was sensitive to hip center location</td>
<td>Crowninshield et al. (1978)</td>
</tr>
<tr>
<td></td>
<td>31 MGs, 7 DoFs</td>
<td>Minimize sum of muscle forces and also mechanico-chemical power output of muscles (a function of muscle rest length, endpoint velocity and zero force velocity)</td>
<td>EMG, predicted muscle forces must not exceed physiological muscle stress capability, muscle force gait patterns should not display excessive inter-subject variation</td>
<td>Biggest source of error in present gait models was depicted as the incomplete information on the physiologic function and role of individual muscles during the gait cycle</td>
<td>Patriarco et al. (1981)</td>
</tr>
<tr>
<td></td>
<td>47 MGs, 3 Joints</td>
<td>Minimize sum of nth power of muscle stresses (n = 1, 2, 3, 4 and 100)</td>
<td>EMG</td>
<td></td>
<td>Crowninshield and Brand (1981)</td>
</tr>
<tr>
<td></td>
<td>42 MGs, 6 DoFs</td>
<td>Minimize sum of muscle forces</td>
<td>Comparison of model outputs to previous literature. Sensitivity analysis of muscle and joint forces to changes in muscle origin/insertion coordinates</td>
<td>Geometrical changes or errors in hip muscle idealization had a greater influence on muscle forces than joint forces at the hip</td>
<td>Rohrle et al. (1984)</td>
</tr>
<tr>
<td></td>
<td>47 MGs, 3 Joints</td>
<td>Maximize endurance by minimizing sum of muscle stresses cubed</td>
<td>Comparison of three solution sets of muscle forces and hip joint forces</td>
<td>Illustrated the sensitivity of muscle force predictions on physiologic cross-sectional area of muscles</td>
<td>Brand et al. (1986)</td>
</tr>
<tr>
<td></td>
<td>7 MGs, 3 Joints</td>
<td>Minimize sum of muscle forces/muscle forces squared/muscle stresses/ligament forces/contact forces/instantaneous muscle power</td>
<td>EMG</td>
<td>All objectives except the principal of minimum total ligament force afforded successful predictions of muscle activation. All minimization principles failed to accurately predict antagonistic activity at the hip and knee</td>
<td>Collins (1995)</td>
</tr>
<tr>
<td></td>
<td>47 MGs, 3 Joints</td>
<td>Minimize sum of muscle forces/muscle forces squared/muscle stresses cubed</td>
<td>EMG</td>
<td>Internal loading was found to depend on the description of joint kinematics. Muscle stresses squared combined with less constrained joints predicted synergistic and antagonist muscle activities</td>
<td>Glitsch and Baumann (1997)</td>
</tr>
<tr>
<td></td>
<td>47 MGs, 3 Joints (maximize endurance)</td>
<td>Minimize sum of muscle stresses cubed</td>
<td>Comparison of model predicted femoral head contact forces to previously reported direct outputs from an instrumented prosthesis</td>
<td>Provided information on the magnitudes and directions of pelvic muscle forces and acetabular contact forces during normal gait</td>
<td>Pedersen et al. (1997)</td>
</tr>
<tr>
<td>Activity</td>
<td>Muscles</td>
<td>Objective</td>
<td>Sensor/Other Consideration</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>----------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>Knee flexion</td>
<td>3 MGs 1 DoF</td>
<td>Minimize sum of muscle forces/muscle stresses</td>
<td>None</td>
<td>Predicted load sharing depended on the choice of decision variable in the objective. When forces were used, activation preference for muscles with large moment arms observed. When stresses were used, activation preference for muscles with product of large moment arms and cross-sectional area were noticed. Linear load sharing criteria predicted orderly muscle recruitment that might apply to onset of muscle action. Non-linear criteria predicted synergistic muscle action.</td>
<td>Dul et al. (1984)</td>
</tr>
<tr>
<td></td>
<td>10 MGs 3 DoF</td>
<td>Minimize sum of muscle forces/joint moments/muscle stresses cubed/muscle activations</td>
<td>EMG</td>
<td>All (linear, non-linear and physiological) optimization criteria predicted antagonistic muscle contraction. Methodology was found to be more sensitive to kinematic information than selection of particular optimization criteria.</td>
<td>Li et al. (1999)</td>
</tr>
<tr>
<td></td>
<td>4 MGs 1 DoF</td>
<td>Minimize sum of muscle activations squared/shifted muscle activations squared</td>
<td>None</td>
<td>A parameter was subtracted from muscle activation to represent stability of the joint and to control antagonistic activity. Co-contraction was predicted through the use of this shift parameter.</td>
<td>Forster et al. (2004)</td>
</tr>
<tr>
<td>Running</td>
<td>47 MGs 3 Joints</td>
<td>Minimize sum of muscle forces/forces squared/muscle stresses/stresses squared/stresses cubed</td>
<td>EMG</td>
<td>Internal loading was found to depend on the description of joint kinematics; muscle stress squared combined with more flexible joints predicted synergistic and antagonistic activities.</td>
<td>Glitsch and Baumann (1997)</td>
</tr>
<tr>
<td>Clenching</td>
<td>26 MGs 6 DoFs</td>
<td>Minimize sum of muscle forces/joint reaction forces</td>
<td>Muscle, joint and bite forces were compared to literature</td>
<td>Minimization of sum of muscle forces were consistent with in vivo observations. Magnitude of bite forces were calculated for a range of bite point locations and bite force directions. The magnitude of maximum possible bite force was found to depend on bite location and direction.</td>
<td>Osborn and Baragar (1985)</td>
</tr>
<tr>
<td></td>
<td>16 MGs 6 DoFs</td>
<td>Minimize relative activity of the most active muscle</td>
<td>Previously published literature</td>
<td></td>
<td>Koolstra et al. (1988)</td>
</tr>
<tr>
<td></td>
<td>10 MGs 1 to 3 DoFs</td>
<td>Minimize joint loads/sum of muscle forces/forces squared/muscle stresses/stresses squared</td>
<td>Ability to predict joint morphologies and muscle outputs with reasonable trends</td>
<td>Minimization of joint load appeared to be more important for morphological development of the temporomandibular joint.</td>
<td>Trainor et al. (1995)</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Activity Model Objective Validation Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinal compression 22 MGs Minimize maximum muscle contraction stress required to satisfy equilibrium EMG</td>
<td>Prediction of muscle forces and lumbar spine loads during isometric tasks imposing extension, bending and twisting; isometric heavy exertions; and level walking</td>
</tr>
<tr>
<td>4 MGs 2 DoFs Minimize maximum muscle intensity/spinal compression force subject to minimized maximum muscle intensity None</td>
<td>A multi-objective, double linear programming approach was proposed as an alternative scheme to solve for muscle forces. In the later study co-contraction was modeled as an incremental increase in the lower bounds of muscle forces</td>
</tr>
<tr>
<td>10 MGs 2 DoFs Minimize sum of cubed muscle stresses/squared muscle stresses/minimum stress-compression/eigenvector synergy EMG</td>
<td>Predictions minimizing muscle stresses cubed were closest to EMG. Muscular contributions to spinal compression force were highly dependent on objective function selection</td>
</tr>
<tr>
<td>10 MGs 3 DoFs Minimize maximum muscle stress and spinal compression (double linear programming)/sum of muscle stresses cubed EMG</td>
<td>Inability of both models to reliably predict measured EMG activity suggested that neither could adequately represent the neural mechanism responsible for generating muscle activation patterns for tested loading conditions (maximum voluntary contractions of trunk flexion and twisting)</td>
</tr>
<tr>
<td>10 MGs 3 DoFs Minimize intensity-compression/muscle intensity cubed None</td>
<td>Influence of objective function selection on peak spinal compression force during mildly asymmetric lifting tasks was found to be minimal</td>
</tr>
<tr>
<td>11 MGs 2 Segments Minimize sum of muscle stresses cubed Comparison against EMG literature</td>
<td>Combined finite element analysis with optimization protocol. Illustrated the importance of muscular loading on the stability and stress distribution of the lumbar spine. Latter study extended similar methodology to 66 muscles spanning 5 lumbar joints to investigate spinal loading during static sagittal plane lifting</td>
</tr>
<tr>
<td>10 MGs 3 DoFs Minimize muscle intensity (force/area) cubed Sensitivity analysis, EMG</td>
<td>Various load combinations of flexion and bending. Reasonable variations of muscle lines of actions (due to anatomical differences and modeling approaches) affected model predictions, particularly shear loading on the spine. Latter study explored Artificial Neural Network classification of muscle recruitment against optimization results</td>
</tr>
<tr>
<td>MGs</td>
<td>DoFs</td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>30</td>
<td>6</td>
</tr>
<tr>
<td>46</td>
<td>6</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>180</td>
<td>6</td>
</tr>
<tr>
<td>52</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
</tr>
</tbody>
</table>

*Continued on next page*
<table>
<thead>
<tr>
<th>Activity</th>
<th>Model</th>
<th>Objective</th>
<th>Validation</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wrist</td>
<td>4 MGs</td>
<td>Minimize sum of muscle forces</td>
<td>Previously published EMG</td>
<td>One of the earlier articles using linear programming principles</td>
<td>Penrod et al. (1974)</td>
</tr>
<tr>
<td>Elbow movements</td>
<td>9 MGs</td>
<td>Minimize upper bound for muscle stresses/sum of muscle forces/muscle stress squared/muscle stresses squared</td>
<td>Previously published EMG data</td>
<td>All objectives predicted the same number of active muscles. Minimizing upper bound of muscle stresses was proposed as a linear programming problem that considered individual muscle effort</td>
<td>An et al. (1984a)</td>
</tr>
<tr>
<td></td>
<td>2 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 MGs</td>
<td>Minimize sum of $n$th power of muscle forces/muscle stresses/relative muscle forces ($n = 1, 2, 3, 10, and 100$)</td>
<td>Comparisons against a validated muscle model driven by maximal muscle activation</td>
<td>Maximally loaded elbow flexion was investigated. Estimated muscle forces were not physiologically realistic when compared to predictions of muscle model. Constant relative force for all muscles (when using relative muscle forces as an objective) was proposed as an alternative for realistic predictions</td>
<td>Challis and Kerwin (1993)</td>
</tr>
<tr>
<td></td>
<td>1 DoF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 MGs</td>
<td>Minimize sum of muscle stress squared/muscle stress cubed/normalized muscle force squared/normalized muscle force cubed</td>
<td>Elbow torque during maximal muscular activity</td>
<td>Submaximal isometric, submaximal dynamic, and maximal dynamic elbow flexions investigated under three separate upper bounds on muscle forces: (i) maximum isometric force, (ii) maximal force as predicted by force–length curve, (iii) maximum muscle force as predicted by force–length and force–velocity curves. Upper bounds were found to be effective on predicted forces</td>
<td>Challis (1997)</td>
</tr>
<tr>
<td></td>
<td>1 DoF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 MGs</td>
<td>Minimize weighted sum of muscle forces squared</td>
<td>None</td>
<td>Introduced Lagrange multipliers to enforce positive and continuous solutions for muscle forces. Also developed a generalized model of the upper limb with 7 DoFs, 30 MGs</td>
<td>Raikova (1992)</td>
</tr>
<tr>
<td></td>
<td>1 DoF</td>
<td></td>
<td></td>
<td>Muscle lever arms (due to different ways of muscle force modeling) exerted a great influence of muscle force predictions. Joint reaction force and muscle forces were not influenced by objective function. Weighting constants for agonistic antagonistic groups were required to have different signs and their magnitude depended on the direction of the net external joint moment. Optimization criterion was formulated for non-negative muscle forces and continuity</td>
<td>Raikova (1996)</td>
</tr>
<tr>
<td></td>
<td>5 MGs</td>
<td>Minimize sum of $n$th power of weighted muscle forces ($n = 2–10$ with a variety of weight constants)</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree of Freedom</td>
<td>Objective Function</td>
<td>Sensitivity Method</td>
<td>Literature Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------</td>
<td>--------------------</td>
<td>----------------------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 5 MGs 1 DoF       | Minimize sum of \(n\)th power (2,3) of weighted muscle forces/muscle activations/\(n\)th power (2) of muscle stress/\(n\)th power (2) of normalized muscle force/weighted combination of error between desired joint torque and torque obtained by summing the individual moments of all motor unit twitches, muscle activation and total muscle force | None | 2,3
| 5 MGs 1 DoF       | Minimize a weighted combination of error between desired joint moment and moment obtained by summing the individual moments of all motor unit twitches, muscle activation and total muscle force | EMG | Raikova and Aladjov (2003)
| 6 MGs 3 DoFs      | Minimize sum of muscle stress cubed | Sensitivity analysis | Pierse and Li (2005)
| Arm and shoulder movements | Minimize muscle stresses squared | Published EMG and shoulder loading data | Karlsson and Peterson (1992)
| 19 MGs 6 DoFs     | Minimization of muscle forces squared/normalized muscle forces squared/maximum muscle stress | EMG | van der Helm (1994)
| 20 MGs 7 DoFs     | Minimize weighted sum of muscle forces squared/cubed | EMG | Happee (1994)

2-3: included calculation of force from a muscle model including force-length, force-velocity relationships. Maximum muscle force used for normalization is a function of force-length, force-velocity relationship in 4. Muscle models incorporated the distribution of motor unit twitches (fast, intermediate, slow). 5 incorporated individual motor units. Hill-type models were found to be suitable for calculating maximum possible forces that can be used as weighting factors. Incorporation of individual function of motor units was promising to explore motor control including co-contraction and the role of fast and slow twitches. Explored learning fast elbow flexions. Muscle models incorporated the distribution of motor unit twitches (fast, intermediate, slow). Objective function incorporated individual motor units.
<table>
<thead>
<tr>
<th>Activity</th>
<th>Model</th>
<th>Objective</th>
<th>Validation</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 MGs 7 DoFs</td>
<td>Minimize metabolic cost (muscular energy consumption)</td>
<td>EMG</td>
<td>Maximal fast goal-directed arm movements in the sagittal plane. Direction of glenohumeral joint was constrained for joint stability. Muscle dynamics was implemented in the form of dynamic boundary constraints</td>
<td>Happee and van der Helm (1995)</td>
<td></td>
</tr>
<tr>
<td>30 MGs 8 DoFs</td>
<td>Minimize sum of muscle forces squared</td>
<td>EMG</td>
<td>Load sharing patterns between shoulder muscles during isometric flexion tasks</td>
<td>Nieminen et al. (1995a)</td>
<td></td>
</tr>
<tr>
<td>30 MGs 9 DoFs</td>
<td>Maximize the endurance time of an activity combined with minimization of muscle forces squared</td>
<td>EMG</td>
<td>Investigated fatiguing static contractions. Model predicted order of fatigue corresponded to EMG signals</td>
<td>Nieminen et al. (1995b)</td>
<td></td>
</tr>
<tr>
<td>30 MGs 8 DoFs</td>
<td>Minimize ratio of current muscle stress and time-dependent maximum muscle stress</td>
<td>None</td>
<td>Novel muscular synergy principle for computing shoulder muscle load sharing. Studied relation between muscle stiffness requirement and muscle co-contraction level</td>
<td>Niemi et al. (1996)</td>
<td></td>
</tr>
<tr>
<td>16 MGs 2 Joints</td>
<td>Minimize sum of muscle forces squared/muscle stresses squared/muscle stresses cubed/normalized muscle force cubed/fatigue</td>
<td>EMG</td>
<td>Isometric joint loading of the arm. Muscle coordination patterns highly depended on the number of balanced DoFs at the elbow. Influence of cost function on results were minimal. Cost functions were adequate to represent actual muscle activity at the wrist but not at the elbow</td>
<td>Buchanan and Shreeve (1996)</td>
<td></td>
</tr>
<tr>
<td>20 MGs 7 DoFs</td>
<td>Maximize hand force in a given direction (linear programming)</td>
<td>Comparison against measured maximum pull strengths</td>
<td>Isometric arm loading during pulling at maximal strength. An arm strength prediction model based on individual muscle strength was provided</td>
<td>Hughes et al. (1999)</td>
<td></td>
</tr>
<tr>
<td>13 MGs 3 DoFs</td>
<td>Minimize muscle stress squared</td>
<td>Previously published muscle forces</td>
<td>Possible to identify muscular function in relation to stages of wheelchair propulsion</td>
<td>Lin et al. (2004)</td>
<td></td>
</tr>
<tr>
<td>31 MGs Not specified</td>
<td>Minimize muscle stress</td>
<td>None</td>
<td>Investigated the influence of tetraplegia and paraplegia on glenoumeral joint and muscle forces while using wheelchairs. Tetraplegic subjects had significantly higher joint contact forces</td>
<td>van Drongelen et al. (2005)</td>
<td></td>
</tr>
</tbody>
</table>

MG: muscle group; DoF: degree of freedom. For studies not reporting number of DoFs, total number of joints or segments is shown.
Muscle force patterns are typically compared to EMG activity patterns to validate the results of the static optimization approach (Table 1). In many cases, similarities have been noted but controversy existed while predicting co-contraction of muscles (Herzog and Binding, 1992; Herzog and Binding, 1993; Herzog and Leonard, 1991; Hughes et al., 1995). Minimizing the sum of cubed muscle stresses as proposed by Crowninshield and Brand (1981) to maximize endurance has been widely accepted for lower extremity analysis and used to predict muscle forces during walking regardless its formulation as a nonlinear optimization problem (Table 1). For the investigation of muscular loading of the upper extremity, minimizing sum of squared muscle forces seems to be more common (Table 1). A variety of performance criteria based on muscle forces has also been tested and compared (Collins, 1995; Li et al., 1999). The computational cost of the technique is small, which allowed sensitivity analyses to be performed in almost all studies (Table 1).

3.2. Forward dynamics assisted data tracking

Another approach that exploits gait data to estimate muscle forces is forward dynamics assisted data tracking. Rather than solving the inverse dynamics problem (Eq. (6)), an initial set of muscle activations are fed into a forward dynamics model of the musculoskeletal system (Eq. (5)). The solution is compared against experimental data and the process is iterated by updating the muscle activations that best reproduce the experimental kinematics and in some cases kinetics, e.g. ground reaction forces, as depicted by the objective $J$ (Fig. 4B, Eq. (8)). Muscle excitations are always limited to unity bound constraints (Eq. (8))

\[
\begin{align*}
\text{minimize } & J(q - q_{\exp}) = J(q) - J(q_{\exp}) \\
\text{subject to } & 0 \leq u \leq 1.
\end{align*}
\]

A typical cost function is $J = ||q - q_{\exp}||$, i.e. least squares fitting of experimental kinematics. Measured external forces may also be included, $J(q - q_{\exp}, E - E_{\exp})$. The technique has been used in a variety of activities and particularly found its applications for high pace movements of sports biomechanics (Table 2). A common use has been to find a set of muscle activations that can reliably reproduce the movement pattern, and subsequently perturb parameters of the optimal solution to explore injury mechanisms (McLean et al., 2004; McLean et al., 2003). This strategy is advantageous due to the more straightforward inclusion of muscle dynamics within the solution when compared to inverse dynamics-based static optimization (Happee, 1994). Although the dynamics of the muscle (activation and force generation properties) might not be influential for low pace movements, muscle force estimation for activities of high performance might benefit from this property of forward dynamics assisted data tracking.

Unlike the inverse dynamics case, kinematic data is incorporated within the forward dynamics model approach in a somewhat weak fashion, allowing muscle force estimations to be less sensitive to measurement errors in kinematic inputs. Regardless of these advantages, however, the technique is computationally involved due to multiple integrations to obtain optimal joint kinematics. Considering these facts, direct application of the forward dynamics assisted data tracking approach to the clinical setting may be difficult, where a rapid output response is often necessary. Recently an efficient technique was proposed to solve the tracking problem (Thelen et al., 2003), which should result in muscle force estimates which are consistent with measured joint torques as well as with known muscle properties.

The forward dynamics methodology is commonly assessed by its performance while tracking experimental data (Table 2). In a majority of the presented cases, the agreement between model predicted movements and kinematic measurements was good. Additional evaluation involved comparison of untracked experimental data (such as pedal forces by Neptune and Hull (1999)) with model predictions. In terms of muscle forces, the validation was solely based on comparing EMG data with estimated muscle activation patterns (Table 2). It is possible that multiple solutions exist to track the same experimental data. Many investigators also explored minimization of a physiological variable, e.g. muscle stress (Yamaguchi and Zajac, 1990), in addition to the tracking error. This multi-objective criterion probably increased the tracking errors in favor of estimating muscular forces based on task objectives. The validity of such an approach and the extent of weighting in between these objectives are not yet known.

3.3. Optimal control strategies

Occasionally the experimental data might be incomplete or the movement related investigations require predictive simulations of the musculoskeletal system in novel situations for which no movement data are available. Under these circumstances, optimal control strategies that use forward dynamics are alternatives to solve for muscle excitations (and forces as by-products) during movements (Pandy, 2001; Pandy et al., 1992). Given an initial set of muscle excitations, system equations are first solved in a forward dynamics fashion (Eq. (5)). Then, the objective $J$ of the movement and task related constraints $h$, e.g. static equilibrium at final time, are calculated (Fig. 4C, Eq. (9)). The objective $J$ can be a function of muscle force and kinematics, can be related to task performance, e.g. maximum height jumping, and it is usually represented in an integral form to introduce dependence on time history. Muscle excitations are always limited to unity bound
<table>
<thead>
<tr>
<th>Activity</th>
<th>Model</th>
<th>Objective</th>
<th>Validation</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking</td>
<td>9 MGs</td>
<td>Minimize tracking error and metabolic energy consumption</td>
<td>Comparison with static optimization, comparisons with EMG estimates from literature</td>
<td>Swing phase only, continuous controls</td>
<td>Davy and Audu (1987)</td>
</tr>
<tr>
<td></td>
<td>3 DoFs</td>
<td>Minimize tracking error and sum of cubed muscle stresses (~muscle fatigue)</td>
<td>Kinematics, GRF, EMG against normative data in literature</td>
<td>Exploration of normal gait simulations as a baseline for functional neuromuscular stimulation</td>
<td>Yamaguchi and Zajac (1990)</td>
</tr>
<tr>
<td></td>
<td>10 MGs</td>
<td>Minimize kinematics and kinetics tracking error</td>
<td>Kinematics and kinetics, EMG</td>
<td>Followed by induced acceleration analysis to quantify individual contributions of ankle plantar flexors to support, forward progression and swing initiation, controls modeled as a block pattern (onset, duration and magnitude)</td>
<td>Neptune et al. (2001)</td>
</tr>
<tr>
<td>Cycling</td>
<td>15 MGs</td>
<td>Minimize tracking error in pedal forces$^1$/crank torque$^2$/joint torques$^3$/crank torque and pedal angle$^4$/2 and 3$^5$/1, 3 and 4$^6$/1, 3, 4 and timing of muscles$^7$</td>
<td>Kinematics, kinetics, EMG</td>
<td>Performance criteria of tracking all variables$^7$ provided the best agreement, controls modeled as block pattern with duration and magnitude</td>
<td>Neptune and Hull (1998)</td>
</tr>
<tr>
<td></td>
<td>2 DoFs</td>
<td>Minimize tracking error in kinematic and kinetic data and timing of muscles</td>
<td>Pedal forces and crank torque</td>
<td>Preferred cycling rate selection in endurance cycling based on neuromuscular quantities (e.g. muscle stress, endurance), controls modeled as block pattern with duration and magnitude</td>
<td>Neptune and Hull (1999)</td>
</tr>
<tr>
<td></td>
<td>9 MGs</td>
<td>Minimize tracking error in pedal forces, crank torque, joint torques</td>
<td>Pedal forces</td>
<td>Comparison of optimization algorithms, controls modeled as block pattern with duration and magnitude</td>
<td>Neptune (1999)</td>
</tr>
<tr>
<td>Jumping</td>
<td>9 MGs</td>
<td>Minimize tracking error and neuromuscular values of muscle excitations</td>
<td>Kinematics, EMG</td>
<td>One-legged jump; controls approximated by polynomials</td>
<td>Spagele et al. (1999)</td>
</tr>
<tr>
<td>Running</td>
<td>14 MGs</td>
<td>Minimize tracking error in segment kinematics and GRF</td>
<td>Joint kinematics, GRF, EMG</td>
<td>Stance phase of heel–toe running, controls modeled as block pattern with duration and magnitude</td>
<td>Neptune et al. (2000)</td>
</tr>
<tr>
<td>Side stepping</td>
<td>11 MGs</td>
<td>Minimize tracking error in kinematics and GRF</td>
<td>Joint angles, vertical GRF, muscle activation comparisons with literature</td>
<td>Prediction of knee joint loading; perturbation to simulation inputs to model subject variability and evaluate potentially hazardous knee joint loading, linear interpolation of controls</td>
<td>McLean et al. (2003)</td>
</tr>
<tr>
<td>Landing in skiing</td>
<td>8 MGs</td>
<td>Minimize tracking error in kinematics</td>
<td>Joint angles</td>
<td>Injury simulation by muscular over stimulation with respect to optimal solution, constant controls during the entire movement</td>
<td>Gerritsen et al. (1996)</td>
</tr>
</tbody>
</table>

Muscle forces are by products of the solution technique. MG: muscle group; DoF: degree of freedom.
constraints (Eq. (9)). The process is iterated until an optimal set of muscle excitation patterns is found that minimizes the objective and satisfies the constraints (Eq. (9))

\[
\text{minimize} \quad J(F_{MT}, q) = \sum h(q) = 0.
\]

Optimal control approaches have been used to investigate muscular function during the activities of daily living such as walking and sit-to-stand, under more physically demanding tasks like jumping and running and for goal-directed movements of the upper limb (Table 3). The technique allows for changes in motion and adaptations at the muscular control level following alterations in the system. This major advantage can lead to predictive simulations to assess changes in control of muscles and muscle forces as a result of therapeutic interventions, surgery and rehabilitation (e.g., plantar fasciotomy by Erdemir and Piazza (2004)).

Muscle activations and movements predicted by the optimal control strategies are usually evaluated by comparisons to joint kinematics data, ground reaction forces and EMG data (Pandy, 2001). In most of the cases, the model predictions are qualitatively in agreement with measured data. However, the selection of an objective function can still be controversial; the criterion is clear for movements that aim for optimal performance (e.g., maximal height jumping) but for other activities (that rely on physiological function) such as walking at different speeds and non-ballistic movements, this selection relies on the investigators’ preference. It is possible that different objective functions lead to similar movement patterns and muscle forces (Pandy et al., 1995). Testing multiple criteria, however, is not always feasible, particularly due to excessive computation time. Computational complexity and implementation difficulties also prohibit the routine use of this technique in clinical settings and limit its use to research environments.

3.4. Alternative strategies

Inaccuracies of inverse dynamics analysis and high computational cost associated with multiple forward dynamics simulations have directed many investigators to search for alternative strategies to estimate muscle forces. This section highlights some of these studies which preferred including EMG data into calculations rather than using the information solely for validation purposes, as well as studies that focused on combining methodologies or fine-tuning algorithms to speed up the solution process.

Using prescribed muscle activations (e.g. maximal activation), a single forward dynamics simulation provides the joint movements and muscle forces. Such an approach was particularly common to investigate jaw mechanics (Koolstra and van Eijden, 1997; Koolstra and van Eijden, 2005; Langenbach and Hannam, 1999). Muscle activations can be prescribed in an educated manner by directly incorporating EMG data into an EMG-driven forward dynamics model, as pioneered by Hof and van den Berg (1981). The approach was also used to investigate elbow motion (Koo and Mak, 2005) and knee movements during the swing phase of walking (Piazza and Delp, 1996) and a step-up task (Piazza and Delp, 2001). The analysis is commonly performed by prescribing the trajectory of some joint angles, e.g. hip and ankle, and predict the remaining from the EMG-driven forward dynamics, e.g. knee movements (Piazza and Delp, 1996; Piazza and Delp, 2001). It is likely that the muscle forces predicted via such approaches suffered from inaccuracies of the muscle model parameters and the processing of the EMG.

EMG data have also been used in combination with the inverse dynamics approach to estimate muscle forces across a series of joints. This process usually involves the calibration of musculoskeletal models and adjusting of muscular gains by minimizing the difference between measured joint torques and those calculated by the EMG-driven model (e.g., Amarantini and Martin (2004)). The methodology was particularly popular to estimate muscle and joint forces during spinal loading (Cromwell et al., 1989; Cholewicki and McGill, 1994; Cholewicki et al., 1995; Cholewicki and McGill, 1996; Granata and Marras, 1995; Nussbaum and Chaffin, 1998; Thelen et al., 1994; van Dieen et al., 2003; Marras et al., 1999). It has also been used for muscle force prediction during arm movements (Manal et al., 2002; Langenderfer et al., 2005; Soechting and Flanders, 1997), isometric tasks of the elbow (Manal and Buchanan, 2003; Wang and Buchanan, 2002) and the wrist (Buchanan et al., 1993). Utilization of the method for the lower extremity included walking (Amarantini and Martin, 2004) and running (Lloyd and Besier, 2003). These studies usually required additional experimentation to obtain the relationship between isometric EMG and joint torque, e.g. “isometric calibration” by maximal isometric effort testing, isometric knee flexion and extension contractions at 20–80% of maximum voluntary contraction with increments of 20% (Amarantini and Martin, 2004). The influence of EMG processing on the estimation of joint moment under dynamic conditions was addressed in detail (Amarantini and Martin, 2004), support for linear EMG to muscle force processing was provided (Raschke and Chaffin, 1996), and novel approaches including processing using neural networks were also proposed (Wang and Buchanana, 2002). Gagnon et al. (2001) compared the ability of EMG-based methods and inverse dynamics-based static optimization to predict spinal loads and trunk muscle forces. Although spinal compression was similar among the approaches, only EMG-based methods detected individual trunk muscle strategies.

Static optimization has also been implemented with forward dynamics. Koolstra and van Eijden (2001) predicted jaw trajectory and muscle forces during goal-directed movements of the jaw, by finding a rest force that maximizes the movement to the desired position at each step.
<table>
<thead>
<tr>
<th>Activity</th>
<th>Model</th>
<th>Objective</th>
<th>Constraints</th>
<th>Validation</th>
<th>Notes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking</td>
<td>54 MGs</td>
<td>Minimization of metabolic energy expenditure per unit distance traveled</td>
<td>Bilateral symmetry, fixed final time, temporal symmetry</td>
<td>Kinematics, GRF, EMG</td>
<td>Linear interpolation of controls</td>
<td>Anderson and Pandy (2001b)</td>
</tr>
<tr>
<td></td>
<td>23 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jumping</td>
<td>8 MGs</td>
<td>Maximum height</td>
<td>Bilateral symmetry</td>
<td>Jump height, lift-off time, qualitative comparisons</td>
<td>Bang-bang (on–off) controls</td>
<td>Pandy et al. (1990)</td>
</tr>
<tr>
<td></td>
<td>4 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8 MGs</td>
<td>Maximum height</td>
<td>Bilateral symmetry</td>
<td>Kinematics, GRF, EMG</td>
<td>Bang-bang (on–off) controls</td>
<td>Pandy and Zajac (1991)</td>
</tr>
<tr>
<td></td>
<td>4 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8 MGs</td>
<td>Maximum height</td>
<td>Bilateral symmetry</td>
<td>Kinematics, GRF, EMG</td>
<td>Linear interpolation of controls</td>
<td>Pandy et al. (1992)</td>
</tr>
<tr>
<td></td>
<td>4 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>54 MGs</td>
<td>Maximum height</td>
<td>Bilateral symmetry</td>
<td>Kinematics, GRF, EMG</td>
<td>Linear interpolation of controls</td>
<td>Anderson and Pandy (1999)</td>
</tr>
<tr>
<td></td>
<td>23 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6 MGs</td>
<td>Maximum height</td>
<td>Bilateral symmetry</td>
<td>Joint kinematics</td>
<td>Controls were defined by time onset of maximal activation</td>
<td>Bobbert (2001)</td>
</tr>
<tr>
<td></td>
<td>3 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20 DoFs</td>
<td></td>
<td>Bilateral symmetry</td>
<td>None</td>
<td>Controls modeled as a series of step function with constant duration. Investigated counter movement jump</td>
<td>Nagano et al. (2005)</td>
</tr>
<tr>
<td></td>
<td>32 MGs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cycling</td>
<td>15 MGs</td>
<td>Maximize crank progress (speed)</td>
<td>Bilateral symmetry</td>
<td>Crank kinematics, crank forces, EMG</td>
<td>Bang-bang (on–off controls)</td>
<td>Raasch et al. (1997)</td>
</tr>
<tr>
<td></td>
<td>3 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15 MGs</td>
<td>Minimize summed integrated muscle activation and average endurance</td>
<td>Bilateral symmetry, average pedaling rate of 90 rpm</td>
<td>Pedal forces and crank torque</td>
<td>Controls modeled as block pattern with duration and magnitude</td>
<td>Neptune and Hull (1999)</td>
</tr>
<tr>
<td></td>
<td>2 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posture</td>
<td>10 MGs</td>
<td>Minimize sum of squared derivatives of adimensional muscle forces (~minimize jerk or maximize smoothness)</td>
<td>Final limb angles and angular velocities set to zero</td>
<td>None</td>
<td>Linear interpolation of controls (small time steps)</td>
<td>Menegaldo et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>3 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising on toes</td>
<td>6 MGs</td>
<td>Minimize squared muscle stress</td>
<td>Bilateral symmetry, fixed final time, static equilibrium at final time</td>
<td>Pelvis kinematics, GRF, EMG</td>
<td>Linear interpolation of controls</td>
<td>Erdemir and Piazza (2004)</td>
</tr>
<tr>
<td></td>
<td>6 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rising from a chair</td>
<td>8 MGs</td>
<td>Minimize squared muscle stress (normalized muscle force)/minimize squared time derivative of normalized muscle force/ combination</td>
<td>Bilateral symmetry, static equilibrium at final time</td>
<td>Vertical GRF and seat forces, EMG</td>
<td>Linear interpolation of controls; better performance of combined criterion</td>
<td>Pandy et al. (1995)</td>
</tr>
<tr>
<td></td>
<td>3 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kicking</td>
<td>5 MGs</td>
<td>Minimum time kicking</td>
<td>Constant final hip and knee angle and zero knee angular velocity</td>
<td>Comparison of results from different muscle models</td>
<td>Continuous controls</td>
<td>Audu and Davy (1985)</td>
</tr>
<tr>
<td></td>
<td>2 DoFs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
of forward dynamics integration. Formulations of static optimization coupled with forward dynamics has also been provided to minimize tracking error of movements (Yamaguchi et al., 1995; Thelen et al., 2003). In these protocols, a single forward dynamics simulation is performed, thereby avoiding the time consuming process associated with standard forward dynamics assisted data tracking. Yamaguchi et al. (1995) calculated muscle forces during arm movements by minimizing the sum of the cubed muscle stresses (Crowninshield and Brand, 1981) at each instant in time, while achieving muscle-induced accelerations that produced a desired trajectory. Thelen et al. (2003) implemented static optimization into a feedback control system to minimize tracking errors in joint kinematics during cycling by minimizing the sum of the squared muscle activations at each step of the forward dynamic solution. Sensitivity of results on the selection of objective function while using these cost-effective tracking algorithms has not been addressed yet.

3.5. Validation

Studies of muscle force predictions usually compare muscle loading or activation patterns against EMG data as an estimate of validity (Tables 1–3). Although evaluating the temporal characteristics and intensity of muscle firing during a movement is useful, such comparisons cannot verify the magnitude of the calculated muscle force. Fortunately, alternative and more advanced analyses exist, which incorporate the quantification of muscle force sensitivity on modeling parameters and comparisons of muscle forces against direct measurements of tendon loading.

Some critical model parameters are associated with the force generation properties of the muscle. Maximum isometric force and physiological cross-sectional area of muscles have been shown to affect the magnitude of muscle force estimates, particularly during inverse dynamics-based static optimization (Brand et al., 1986). Force–length, force–velocity properties and activation dynamics can have significant effects when using forward dynamics assisted data tracking or optimal control solutions (Audu and Davy, 1985). Also important, as noted earlier, are the origin, insertion, and path of muscles relative to the skeleton. These variables determine the moment arms for each muscle; an increased moment arm would result in a decreased muscle force estimate while producing the same moment at a given joint (Raikova and Prilutsky, 2001).

Descriptions of joint mechanics also have impacts on the distribution of muscular forces. A variety of joint types are available for modeling, from simple hinge joints to the three-dimensional spherical joints. The joint type defines the constraints of movements and has been found to influence muscle force calculations (Buchanan and Shreeve, 1996; Glitsch and Baumann, 1997). For example, in a spherical joint, all joint torques need to be balanced by muscular moments, whereas for a hinge joint, out of plane torques are balanced by reaction torques at the joint.
Consequently, muscle forces can be overestimated when the physiological joint constraints are not imposed, e.g. modeling the ankle joint as a spherical joint (Burdett, 1982). Related to kinematics, a two-dimensional musculoskeletal model will give different results than a three-dimensional model as a result of enforcing joint movements, which are naturally three-dimensional, to be in a plane (Glitsch and Baumann, 1997). A recent study also illustrated the dependence of muscle force estimations on coordinate system definitions (Pierce and Li, 2005).

Among simulation parameters, selection of the objective function might have an effect on the prediction of muscle forces. Multiple physiological criteria have been tested while calculating muscular loading using inverse dynamics-based static optimization (Table 1). Although a wide variety of these objective functions resulted in similar muscle forces, some seemed to work better depending on the movement pattern under investigation (e.g. Crowninshield and Brand (1981)), and when predicting co-activation of muscles (Forster et al., 2004; Li et al., 1999; Hughes et al., 1995). Objective function definition is clear for forward dynamics assisted data tracking. Further, models incorporating tracking of both kinematic and kinetic variables (e.g. external forces) were less erroneous than those including kinematic variables only (Neptune and Hull, 1998). Identification of a performance criterion is critical in optimal control simulations as a result of the generated movement depending on that criterion as well as the constraints of the task. The movement patterns are supplied in inverse dynamics-based static optimization and forward dynamics assisted data tracking. Pandy et al. (1995) have shown that similar movement patterns can be obtained using optimal control simulations with different objective functions while investigating non-ballistic activities, but the muscle activation patterns might be different.

Accuracy of the experimental data used within the model is paramount for accurate muscle force estimations. Errors in joint kinematics and calculated joint torques for example, are known to significantly alter modeled muscle force magnitudes, particularly when the calculations are done using inverse dynamics and static optimization (Li et al., 1999; Glitsch and Baumann, 1997). Anderson and Pandy (2001a) illustrated that it is possible to predict similar muscle forces and joint reaction forces for walking using the inverse dynamics-based static optimization approach and the optimal control simulation approach. They used the joint kinematics and torques calculated from the forward dynamics simulation of the optimal control problem as input to the inverse dynamics-based static optimization. The consistency observed in these muscle force predictions suggests that if experimental accuracy can be improved, then resultant muscle forces might not depend on the simulation characteristics.

Direct validation of predicted muscle forces is possible by comparing them against the tendon forces measured experimentally. Typically, direct validations are limited to simple musculoskeletal models, e.g. with one or two degrees of freedom (Binding et al., 2000), and tendon force measurements are performed on animals by surgical implantation of tendon force measurement devices (Herzog and Leonard, 1991; Landjerit et al., 1988). Nonetheless, the results of these studies can be used to assess the validity of objective functions used in inverse dynamics-based static optimization and the load sharing between synergistic muscles (Binding et al., 2000). Validation data for pathological conditions however do not currently exist and unfortunately, direct validation in humans using current in vivo tendon force transducers is limited. The devices are invasive (Ravary et al., 2004; Fleming and Beynnon, 2004), can only be placed on extrinsic tendons (e.g., the Achilles tendon by Komi et al. (1992)) or used in the operation room during surgery (Dennerlein et al., 1998; Dennerlein et al., 1999; Dennerlein, 2005; Schuind et al., 1992), possess measurement errors inherent to transducer design (Ravary et al., 2004), and transducer calibration remains a largely unsolved problem. In large-scale musculoskeletal models, validity of muscle force estimates has been assessed indirectly by comparing measured joint reaction forces against those predicted by modeling. Brand et al. (1994) compared hip forces predicted by a musculoskeletal model to direct measurements from an instrumented hip implant. They reported that muscle forces were apparently overestimated due to the lack of realistic wrapping of muscle paths around the hip joint, making moment arms smaller than they should have been.

4. Induced acceleration analysis

The effect of muscular loading on joint kinematics is clear for uniarticular muscles and for simple movements. However, the influence of individual muscle forces on kinematics may not be identified easily when the movement pattern is complicated and contains musculoskeletal coupling and the involvement of multiple muscles and joints, e.g. walking. Induced acceleration analysis (IAA) provides a platform to establish the link between an isolated change in a muscle force and the corresponding changes in the movement. The induced accelerations of muscle $j$ is defined as the contribution of muscle $j$ to all system accelerations:

$$ IAA_j = \Delta M(q)^{-1}R^{(j)}(q)F_{MTj}, \quad (10) $$

where $R^{(j)}(q)$ is the $j$th column of the muscular moment arm matrix $R(q)$. Zajac and Gordon (1989) showed that the elements of the matrix $M(q)^{-1}R^{(j)}(q)$ in Eq. (10) are non-zero. This implies that each muscle contributes to motion of all joints. This “coupled dynamics” representation can explain some of the counterintuitive functions of biarticular muscles, such as the gastrocnemius functioning as knee extensor for specific conditions (Zajac and Gordon, 1989). Usually, the analysis is based on forward dynamics solutions obtained through data tracking or by optimal control (Anderson and Pandy, 2003; Neptune et al., 2001; Zajac et al., 2003). However, data can also be entered directly.
into Eq. (10) through the measured kinematics \( (\mathbf{q}(t)) \). If the musculoskeletal model includes contact with the ground, special care is needed to avoid incorrect results \( (\text{Neptune et al., 2001; Kepple et al., 1997}) \).

A common IAA variable is the trunk acceleration. Its vertical component provides the contribution of individual muscles to keep the trunk supported during locomotion; the horizontal component gives a measure of contribution to forward progression \( (\text{Neptune et al., 2001}) \). IAA has also been presented in the form of induced position or velocity analysis. In that case, a small time-step integration of Eq. (10) is needed to quantify the influence of an individual muscle on the kinematic variable of interest; for example, knee flexion \( (\text{Anderson et al., 2004}) \) and knee flexion velocity \( (\text{Goldberg et al., 2004}) \). Vertical ground reaction force has also been used to implicitly evaluate the contribution of each muscle to support the body center of mass \( (\text{Anderson and Pandy, 2003}) \).

5. Recommendations for clinical applications

The ability to obtain quantitative estimates of muscle forces during movement has significant clinical potential, which has not yet been realized. Before considering such clinical applications, it is important to balance the potential usefulness of this approach against its limitations.

Clinical estimation of muscle forces can be compared to the more traditional techniques of EMG and inverse dynamics \( (\text{quantification of joint moments}) \). While EMG is a direct measurement of individual muscle activity, the magnitude of the signal is affected by electrode placement and tissue conductivity \( (\text{De Luca, 1997}) \), which makes it difficult to use EMG as an indicator of muscle force when comparing between patients and control subjects, or between testing sessions such as before and after treatment \( (\text{Clark et al., in press}) \). Surface EMG is only applicable to superficial muscles. Intramuscular EMG does not have this limitation but is even less correlated to muscle force, limiting its use to studying the timing of muscle activity. Inverse dynamic analysis is well established but it only provides an estimate of the total joint torques produced by all muscles that cross a joint. Interpretation of joint torques often involves a decision process about which muscles could be responsible for the observed results. This interpretation is typically done for one joint at a time, and can be incorrect unless the clinician has a good understanding of functional anatomy and is experienced in combining the joint torques of the entire limb and possible contributions from each muscle. For instance, if a knee extensor moment is small, this is usually attributed to an abnormally small quadriceps force \( (\text{Berchuck et al., 1990}) \). An equally plausible interpretation, however, is that the hamstrings force is abnormally large. One would have to examine the hip extensor moment to resolve the ambiguity, but even then, this is not straightforward because the hamstrings are not the only muscles contributing to the hip joint torque. Any technique which estimates individual muscle forces, such as those reviewed in this article, automatically includes all these considerations. Therefore, we recommend that estimation of muscle forces should be considered if the clinical question requires an interpretation in terms of individual muscle forces, and when a useful interpretation cannot be obtained from joint torques or EMG. Typically this would be the case when the clinical problem involves co-contraction of antagonistic muscles and a case-control or longitudinal study design where EMG is not applicable. Muscular co-contraction is important in injury prevention and rehabilitation \( (\text{Hurd et al., 2006}) \) and in applications where articular forces must be estimated \( (\text{Crowninshield et al., 1978}) \). On the other hand, if the clinical question can be answered with EMG or joint torques alone, we recommend pursuing that path, rather than introducing the additional complexity and uncertainty of estimating individual muscle forces.

Once a specific clinical need for estimating muscle forces has been established, the clinical potential must be balanced against the many limitations of the methodology. First, there are currently no user-friendly software systems that allow estimation of muscle forces from kinesiological data. Vendors of data acquisition systems typically provide software to evaluate joint angles and joint torques, which still require significant post-processing in order to obtain muscle forces. In the above review, we have pointed out the importance of using a model with the appropriate number of degrees of freedom in order to avoid overestimation of muscle forces. This means that the standard 3D joint torques are not necessarily a good starting point. We refer to Table 1 for guidance on the design of appropriate models and algorithms for the various subsystems in the body. The lack of software means that clinical researchers will likely require help from programmers, but also scientific input into the design of the mechanical models. Second, there is a need for preliminary research to establish and validate a suitable method for estimating muscle forces in each particular application. Previous research \( (\text{Table 1}) \) has quantified sensitivity of muscle force estimations and resulted in qualitative or semi-quantitative \( (\text{EMG-based}) \) validations for specific muscles, specific movements, and mostly in healthy populations. In addition, most optimization algorithms are based on the notion that muscles are coordinated according to a minimal effort principle. In a patient with pain, muscle coordination may be guided more by the desire to avoid mechanical stress on the painful tissue. In a patient with a neurological disorder, muscle coordination may be affected by spasticity or paralysis, which would no longer allow the central nervous system to use a minimal energy principle. Nevertheless, estimates may still be adequate if muscle forces predictions are constrained sufficiently by measured kinematics or joint torques. As discussed in this paper, force estimates may be constrained even more by considering known mechanical properties of muscle and measured EMG signals. Therefore, the validity of muscle force estimation, or at least its sensitivity, must be established for each clinical application with the intention to estimate study-specific error margins.
6. Recommendations for research

This review of literature points towards opportunities for research in three areas: clinical, computational, and experimental.

Clinical research is necessary in order to develop and encourage important applications. This includes development of clinical protocols and their validation, as discussed in the previous section. Once such protocols are established, it would also be important to perform clinical studies aimed at demonstrating that such analyses may lead to improved clinical decision making and ultimately, a better health outcome or lower cost of treatment. Without such research, cost reimbursements for eventual clinical applications will remain problematic.

Development of user-friendly computational tools is important to support clinical applications. These tools might combine existing algorithms and musculoskeletal models (e.g. Table 1) and make them accessible to a broader group of users. Several commercial initiatives are currently being undertaken, such as by Musculographics, Inc. (Chicago, IL), the Anybody Group (Aalborg, Denmark), and Biomechanics Research Group, Inc. (San Clemente, CA) but their use is still limited to research. Clinical applications have not been reported, possibly because the software is still difficult to use and lacks scientific validation.

Research on new algorithms is also warranted, especially in optimal control methods to predict movements from optimization principles without use of measurements. This type of application is important for surgical planning (e.g. tendon transfers) and in prosthetic design. If movement can be predicted, there is an opportunity for a true computer-aided design of the human–machine system. Current movement prediction algorithms are extremely computer-intensive (Anderson and Pandy, 2001a,b) which necessitates parallel computing (Neptune, 1999; Anderson et al., 1995; van Soest and Casius, 2003; Koh et al., 2004). A promising approach lies in collocation methods (Kaplan and Heegaard, 2001), which may solve the movement prediction problem on standard desktop hardware. Computationally effective tools are also evolving for data tracking which combines static optimization with a single forward dynamic simulation (Yamaguchi et al., 1995; Thelen et al., 2003; Thelen and Anderson, 2006). In a mature state, these tools should allow for integration of kinematics, external forces, and EMG to improve the reliability of muscle force estimations.

A third area of computational research is in musculoskeletal modeling. It is not known to which extent it is important to obtain patient-specific model parameters, such as muscle moment arms, skeletal anatomy, and mass properties. Subject-specific models may be especially important in clinical populations with bone deformities or altered muscle properties (Arnold et al., 2000, 2001).

Experimental studies are needed to further establish the validity of muscle force estimations by comparison to direct measurements or EMG. Direct measurements are currently only possible by transducers that are implanted and difficult to calibrate (Ravary et al., 2004; Fleming and Beynnon, 2004). A less invasive technique based on ultrasound is a promising alternative (Pourcelot et al., 2005). With appropriate statistical techniques, such as correlation, uncalibrated transducers and semi-quantitative measurements (e.g. EMG) can be effectively used for validation of model-based muscle force estimates (De Zee et al., in press). There is especially a need for validation in clinical populations in which the minimal effort assumptions may not be valid. On a more fundamental level, such research may be done in animal models.

References


